

# Robustness Analysis of Preconditioned Successive Projection Algorithm for General Form of Separable NMF Problem

Tomohiko Mizutani \*

January 29, 2016

## Abstract

The successive projection algorithm (SPA) has been known to work well for separable non-negative matrix factorization (NMF) problems arising in applications, such as topic extraction from documents and endmember detection in hyperspectral images. One of the reasons is in that the algorithm is robust to noise. Gillis and Vavasis showed in [SIAM J. Optim., 25(1), pp. 677–698, 2015] that a preconditioner can further enhance its noise robustness. The proof rested on the condition that the dimension  $d$  and factorization rank  $r$  in the separable NMF problem coincide with each other. However, it may be unrealistic to expect that the condition holds in separable NMF problems appearing in actual applications; in such problems,  $d$  is usually greater than  $r$ . This paper shows, without the condition  $d = r$ , that the preconditioned SPA is robust to noise.

**Keywords:** nonnegative matrix factorization, separability, successive projection algorithm, robustness to noise, preconditioning

**AMS classification:** 15A23, 15A12, 65F30, 90C25

## 1 Introduction

A  $d$ -by- $m$  nonnegative matrix  $\mathbf{A}$  is said to be *separable* if it has a decomposition of the form

$$\mathbf{A} = \mathbf{F}\mathbf{W} \text{ for } \mathbf{F} \in \mathbb{R}_+^{d \times r} \text{ and } \mathbf{W} = (\mathbf{I}, \mathbf{K})\mathbf{\Pi} \in \mathbb{R}_+^{r \times m} \quad (1)$$

where  $\mathbf{I}$  is an  $r$ -by- $r$  identity matrix,  $\mathbf{K}$  is an  $r$ -by- $(m - r)$  nonnegative matrix, and  $\mathbf{\Pi}$  is an  $m$ -by- $m$  permutation matrix. Here, we call  $\mathbf{F}$  the *basis matrix* of  $\mathbf{A}$  and  $r$  the *factorization rank*. The separable nonnegative matrix factorization problem is stated as follows.

**(Separable NMF Problem)** Let  $\mathbf{A}$  be of the form given in (1). Find an index set  $\mathcal{I}$  with  $r$  elements such that  $\mathbf{A}(\mathcal{I})$  coincides with the basis matrix  $\mathbf{F}$ .

The notation  $\mathbf{A}(\mathcal{I})$  denotes the submatrix of  $\mathbf{A}$  whose column indices are in  $\mathcal{I}$ ; in other words,  $\mathbf{A}(\mathcal{I}) = (\mathbf{a}_i : i \in \mathcal{I})$  for the  $i$ th column vector  $\mathbf{a}_i$  of  $\mathbf{A}$ . We use the abbreviation NMF to refer to nonnegative matrix factorization. The problem above can be thought of as a special case of

---

\*Department of Industrial Engineering and Management, Tokyo Institute of Technology, 2-12-1-W9-69, Ookayama, Meguro, Tokyo, 152-8552, Japan. mizutani.t.ab@m.titech.ac.jp

NMF problem. The NMF problem is intractable, and in fact, was shown to be NP-hard in [13]. The authors of [3] proposed to put an assumption, called separability. The separability assumption turns it into a tractable problem referred to as a separable NMF problem. Although the assumption may restrict the range of applications, it is known that separable NMF problems nonetheless can be used for the purpose of topic extraction from documents [4, 2, 11] and endmember detection in hyperspectral images [7, 8].

Several algorithms have been developed for solving the separable NMF problem. One of our concerns is how robust these algorithms are to noise, since it is reasonable to suppose that the separable matrix contains noise in separable NMF problems arising from the applications mentioned above. We consider an algorithm for solving a separable NMF problem and suppose that the separable matrix contains noise. If the algorithm can identify a matrix close to the basis matrix, we say that it is *robust to noise*.

The successive projection algorithm (SPA) was originally proposed in [1] in the context of chemometrics. Currently, the algorithm and its variants are used for topic modeling, document clustering and hyperspectral image unmixing. Gillis and Vavasis showed in [7] that SPA is robust to noise and presented empirical results suggesting that the algorithm is a promising approach to hyperspectral image unmixing. The theoretical results implied that further improvement in noise robustness can be expected if we can make the condition number of the basis matrix smaller. Hence, they proposed in [8] to use a preconditioning matrix for reducing the condition number of the basis matrix. They showed that the noise robustness of SPA is improved by using a preconditioner. The proof rested on the condition that the dimension  $d$  and factorization rank  $r$  in a separable matrix coincide with each other. However, it may be unrealistic to expect that the condition holds in separable NMF problems derived from actual applications. In such a situation,  $d$  is usually greater than  $r$ . For instance, we shall consider extraction of topics from a collection of newspaper articles. This task can be modeled as a separable NMF problem. In the problem, the dimension  $d$  and factorization rank  $r$  of a separable matrix correspond to the number of articles and topics, respectively. It would be rare that  $d$  is close to  $r$  but usual that  $d$  is greater than  $r$ .

The aim of this paper is to show, without the condition  $d = r$ , that the preconditioned SPA is robust to noise. The statement of our result is in Theorem 3. It can be used as a guide for seeing how robust the algorithm is to noise when handling separable NMF problems derived from actual applications.

The rest of this paper is organized as follows. In Section 2, we review SPA and the preconditioned one and describe the results of the noise robustness obtained in a series of studies by Gillis and Vavasis. After that, we describe our analysis of the preconditioned SPA, comparing our results with those of Gillis and Vavasis. Our analysis is shown in Section 3.

## 1.1 Notation and Terminology

A real matrix is said to be *nonnegative* if all of its elements are nonnegative. Here, we use the symbol  $\mathbb{R}^{d \times m}$  to represent the set of  $d$ -by- $m$  real matrices, and  $\mathbb{R}_+^{d \times m}$  the set of  $d$ -by- $m$  nonnegative matrices. The identity matrix is denoted by  $\mathcal{I}$  and the permutation matrix by  $\Pi$ . The vector of all ones is denoted by  $\mathbf{e}$  and the  $i$ th unit vector by  $\mathbf{e}_i$ . We shall use the capital upper-case letter  $\mathbf{A}$  to denote a matrix. The lower-case letter with subscript  $\mathbf{a}_i$  indicates the  $i$ th column. We denote the transpose by  $\mathbf{A}^\top$ , the rank by  $\text{rank}(\mathbf{A})$  and the matrix norm by  $\|\mathbf{A}\|$ . In particular, the matrix 2-norm and the Frobenius norm are written as  $\|\mathbf{A}\|_2$  and  $\|\mathbf{A}\|_F$ . We use the symbol  $(\mathbf{A}; \mathbf{B})$  for

matrices  $\mathbf{A} \in \mathbb{R}^{d \times m}$  and  $\mathbf{B} \in \mathbb{R}^{d' \times m}$  to represent the matrix

$$\begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix} \in \mathbb{R}^{(d+d') \times m},$$

and the symbol  $\text{diag}(a_1, \dots, a_m)$  for numbers  $a_1, \dots, a_m$  to represent the diagonal matrix

$$\begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_m \end{pmatrix} \in \mathbb{R}^{m \times m}.$$

For two real numbers  $a$  and  $b$ , the symbol  $\min(a, b)$  indicates the smaller value.

## 1.2 Tools from Linear Algebra

Any real and complex matrix has a singular value decomposition (SVD). We will use the SVD of a real matrix in our subsequent discussion. Let  $\mathbf{A} \in \mathbb{R}^{d \times m}$ . The SVD of  $\mathbf{A}$  can be written as

$$\mathbf{A} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^\top. \quad (2)$$

$\mathbf{U}$  and  $\mathbf{V}$  are  $d$ -by- $d$  and  $m$ -by- $m$  orthogonal matrices. In particular, the column vectors of  $\mathbf{U}$  and  $\mathbf{V}$  are called the *left singular vectors* and *right singular vectors* of  $\mathbf{A}$ .  $\mathbf{\Sigma}$  is a  $d$ -by- $m$  diagonal matrix. If  $d \leq m$ , it has the form  $(\text{diag}(\sigma_1, \dots, \sigma_d), \mathbf{0})$  for a  $d$ -by- $(m-d)$  zero matrix  $\mathbf{0}$ ; otherwise,  $(\text{diag}(\sigma_1, \dots, \sigma_m); \mathbf{0})$  for a  $(d-m)$ -by- $m$  zero matrix  $\mathbf{0}$ . Let  $t = \min(d, m)$ . It is known that the diagonal elements  $\sigma_1, \dots, \sigma_t$  are all nonnegative. These elements are called *singular values* of  $\mathbf{A}$ . By changing the order of columns in  $\mathbf{U}$  and  $\mathbf{V}$ , we can arrange the singular values in descending order. Therefore, throughout this paper, we always assume that  $\sigma_1 \geq \dots \geq \sigma_t$  in  $\mathbf{\Sigma}$ . We use the symbols  $\sigma_{\min}(\mathbf{A})$  and  $\sigma_{\max}(\mathbf{A})$  to denote the smallest and largest singular values among them; in other words,  $\sigma_{\min}(\mathbf{A}) = \sigma_t$  and  $\sigma_{\max}(\mathbf{A}) = \sigma_1$ . We define the *condition number* of  $\mathbf{A}$  as  $\sigma_{\max}(\mathbf{A})/\sigma_{\min}(\mathbf{A})$ , and use  $\kappa(\mathbf{A})$  to denote it.

Let  $\mathbf{A}$  be a  $d$ -by- $d$  symmetric positive definite matrix. Due to the positive definiteness,  $\mathbf{A}$  has an eigenvalue decomposition such that  $\mathbf{A} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^\top$  where  $\mathbf{U}$  is a  $d$ -by- $d$  orthogonal matrix, and  $\mathbf{\Lambda}$  is a  $d$ -by- $d$  diagonal matrix with positive elements  $\lambda_1, \dots, \lambda_d$ . We define the square root of  $\mathbf{A}$  as  $\mathbf{U} \mathbf{\Lambda}^{1/2} \mathbf{U}^\top$  where  $\mathbf{\Lambda}^{1/2} = \text{diag}(\lambda_1^{1/2}, \dots, \lambda_d^{1/2})$ , and use  $\mathbf{A}^{1/2}$  to denote it. We use the symbol  $\mathbf{A} \succ \mathbf{0}$  to mean that  $\mathbf{A}$  is positive definite.

## 2 Noise Robustness of the Preconditioned SPA for Separable NMF Problems

This section consists of three subsections. We start by examining separable NMF problems from a geometric point of view and summarize each step of SPA in Algorithm 1. The geometric interpretation of the problems will help us to understand the notion behind SPA. The analysis of the noise robustness of SPA by [7] is described in Theorem 1. Next, assuming the condition  $d = r$ , we intuitively explain why the noise robustness of SPA can be enhanced by using a preconditioning matrix. Then we describe the result for the preconditioned SPA by [8] in Theorem 2. We summarize the algorithm of preconditioned SPA in Algorithm 2, which works without the condition  $d = r$ . Finally, we describe our results on the robustness of Algorithm 2 to noise and compare them with the results of Theorem 2 of [8].

## 2.1 Review of SPA

Separable NMF problems have a geometric interpretation. Let  $\mathbf{A}$  be of the form given in (1). Without loss of generality, we can assume that any column vector  $\mathbf{k}_i$  of  $\mathbf{K}$  satisfies  $\mathbf{e}^\top \mathbf{k}_i = 1$ , since  $\mathbf{A} = \mathbf{F}\mathbf{W} \Leftrightarrow \mathbf{A}\mathbf{D}_1 = \mathbf{F}\mathbf{D}_2\mathbf{D}_2^{-1}\mathbf{W}\mathbf{D}_1$  for nonsingular diagonal matrices  $\mathbf{D}_1$  and  $\mathbf{D}_2$ . In addition, assume that  $\text{rank}(\mathbf{F}) = r$ . Under these assumptions, the convex hull of the column vectors of  $\mathbf{A}$  is an  $(r - 1)$ -dimensional simplex in  $\mathbb{R}^d$ , and the vertex corresponds to each column vector of  $\mathbf{F}$ . Accordingly, we can restate the separable NMF problem as follows; find all vertices of the convex hull of the column vectors of  $\mathbf{A}$ . In [7, 8, 11], we can find a further explanation of the problem.

SPA is designed on the basis of the geometric interpretation of separable NMF problems. The first step finds  $\mathbf{a}_{i^*}$  among the column vectors  $\mathbf{a}_1, \dots, \mathbf{a}_m$  of  $\mathbf{A}$  that maximizes the convex function  $f(\mathbf{x}) = \|\mathbf{x}\|_2^2$  and projects  $\mathbf{a}_1, \dots, \mathbf{a}_m$  onto the orthogonal space to  $\mathbf{a}_{i^*}$ . This procedure is repeated until  $r$  column vectors are found. As pointed out in [7], SPA has a connection to QR factorization with column pivoting by [5]. Algorithm 1 describes each step of SPA. We may see why SPA can find  $\mathbf{F}$  from  $\mathbf{A}$  by recalling the following property; given the set of points in a polytope, including all the vertices, the maximum of a strongly convex function over the set is attained at one of the vertices.

---

### Algorithm 1 SPA

---

**Input:** A  $d$ -by- $m$  real matrix  $\mathbf{A}$  and a positive integer  $r$ .

**Output:** An index set  $\mathcal{I}$ .

---

**1:** Initialize a matrix  $\mathbf{S}$  as  $\mathbf{S} \leftarrow \mathbf{A}$ , and an index set  $\mathcal{I}$  as  $\mathcal{I} \leftarrow \emptyset$ .

**2:** Find an index  $i^*$  such that  $i^* = \arg \max_{i=1, \dots, m} \|\mathbf{s}_i\|_2^2$  for the column vector  $\mathbf{s}_i$  of  $\mathbf{S}$ .

**3:** Set  $\mathbf{t} \leftarrow \mathbf{s}_{i^*}$ . Update  $\mathbf{S}$  as

$$\mathbf{S} \leftarrow \left( \mathbf{I} - \frac{\mathbf{t}\mathbf{t}^\top}{\|\mathbf{t}\|_2^2} \right) \mathbf{S},$$

and  $\mathcal{I}$  as

$$\mathcal{I} \leftarrow \mathcal{I} \cup \{i^*\}.$$

**4:** Go back to step 2 if  $|\mathcal{I}| < r$ ; otherwise, output  $\mathcal{I}$ , and then terminate.

---

Now let us describe the analysis of Algorithm 1 given by Gillis and Vavasis in [7]. We put the following assumption on a  $d$ -by- $m$  real matrix  $\mathbf{A}$ .

**Assumption 1.**  $\mathbf{A}$  can be decomposed into  $\mathbf{A} = \mathbf{F}\mathbf{W}$  for  $\mathbf{F} \in \mathbb{R}^{d \times r}$  and  $\mathbf{W} = (\mathbf{I}, \mathbf{K})\mathbf{\Pi} \in \mathbb{R}_+^{r \times m}$  where  $\mathbf{I}$ ,  $\mathbf{K}$  and  $\mathbf{\Pi}$  are the same as those of (1).  $\mathbf{F}$  and the column vector  $\mathbf{k}_i$  of  $\mathbf{K}$  satisfy the following conditions.

(a)  $\text{rank}(\mathbf{F}) = r$ .

(b)  $\mathbf{e}^\top \mathbf{k}_i \leq 1$  for all  $i = 1, \dots, m - r$ .

Assumption 1 corresponds to that made in [7, 8]. Note that the decomposition of  $\mathbf{A}$  in the assumption is not exactly the same as that of (1).  $\mathbf{F}$  in (1) is a nonnegative matrix but  $\mathbf{F}$  in the assumption is not necessarily a nonnegative one. As we mentioned in the first part of this section, from the relation  $\mathbf{A} = \mathbf{F}\mathbf{W} \Leftrightarrow \mathbf{A}\mathbf{D}_1 = \mathbf{F}\mathbf{D}_2\mathbf{D}_2^{-1}\mathbf{W}\mathbf{D}_1$  for nonsingular diagonal matrices  $\mathbf{D}_1$

and  $\mathbf{D}_2$ , Assumption 1(b) can be assumed without loss of generality. Under Assumption 1, Gillis and Vavasis showed in [7] that Algorithm 1 on the input  $(\mathbf{A}, r)$  returns  $\mathcal{I}$  such that  $\mathbf{A}(\mathcal{I}) = \mathbf{F}$ . Furthermore, they showed that it is robust to noise. Suppose that a separable matrix  $\mathbf{A}$  of (1) contains a noise matrix  $\mathbf{N} \in \mathbb{R}^{d \times m}$  such that

$$\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{N}. \quad (3)$$

We call  $\tilde{\mathbf{A}}$  a *near-separable matrix*, and  $\mathbf{N}$  a *noise matrix*. Their analysis tells us that running Algorithm 1 on the input  $(\tilde{\mathbf{A}}, r)$  returns  $\mathcal{I}$  such that  $\tilde{\mathbf{A}}(\mathcal{I})$  is close to  $\mathbf{F}$  if the size of  $\mathbf{N}$  is small. The formal statement is as follows.

**Theorem 1** (Theorem 3 of [7]). *Let  $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{N}$  for  $\mathbf{A} \in \mathbb{R}^{d \times m}$  and  $\mathbf{N} \in \mathbb{R}^{d \times m}$ . Suppose that  $r \geq 2$  and  $\mathbf{A}$  satisfies Assumption 1. If  $\mathbf{n}_i$  of  $\mathbf{N}$  satisfies  $\|\mathbf{n}_i\|_2 \leq \epsilon$  for all  $i = 1, \dots, m$  with*

$$\epsilon < \min\left(\frac{1}{2\sqrt{r-1}}, \frac{1}{4}\right) \frac{\sigma_{\min}(\mathbf{F})}{1 + 80\kappa(\mathbf{F})^2},$$

*then, Algorithm 1 with the input  $(\tilde{\mathbf{A}}, r)$  returns the output  $\mathcal{I}$  such that there is an order of the elements in  $\mathcal{I}$  satisfying*

$$\|\tilde{\mathbf{a}}_{\mathcal{I}(j)} - \mathbf{f}_j\|_2 \leq (1 + 80\kappa(\mathbf{F})^2)\epsilon$$

*for all  $j = 1, \dots, r$ .*

The notation  $\mathcal{I}(j)$  represents the  $j$ th element of  $\mathcal{I}$  for a set  $\mathcal{I}$  whose elements are arranged in some order. Throughout this paper, we will use the notation  $\mathcal{I}(j)$  to refer to the  $j$ th element in the ordered elements of  $\mathcal{I}$ . In the theorem,  $\tilde{\mathbf{a}}_{\mathcal{I}(j)}$  is the  $\mathcal{I}(j)$ th column vector of  $\tilde{\mathbf{A}}$ , and  $\mathbf{f}_j$  the  $j$ th column vector of  $\mathbf{F}$ . The statement of the above theorem does not completely match that of Theorem 3 of [7]. Let us remark on that.

**Remark 1.** Theorem 3 of [7] is described by using  $L$ ,  $\mu$ , and  $K(\mathbf{F})$  that are not found in the above theorem. Let  $f$  be a strongly convex function. Then,  $L$  corresponds to the Lipschitz constant of  $f$ , and  $\mu$  is the parameter associated with the strong convexity of  $f$ . We have  $L = \mu$  since we consider the case in which  $f(\mathbf{x}) = \|\mathbf{x}\|_2^2$ .  $K(\mathbf{F})$  is defined as  $K(\mathbf{F}) = \max_{j=1, \dots, r} \|\mathbf{f}_j\|_2$ . From the definition, we have  $K(\mathbf{F}) \leq \sigma_{\max}(\mathbf{F})$ . Therefore, the above theorem follows from Theorem 3 of [7].

## 2.2 Preconditioned SPA

Consider a near-separable matrix  $\tilde{\mathbf{A}}$  of (3). Theorem 1 suggests that, if one restricts the condition number of the basis matrix  $\mathbf{F}$  to be close to one, we may expect that the allowed range size of  $\|\mathbf{n}_i\|_2$  increases and the difference between  $\mathbf{F}$  and  $\tilde{\mathbf{A}}(\mathcal{I})$  decreases. Assume that  $\mathbf{A}$  of  $\tilde{\mathbf{A}}$  satisfies Assumption 1. Let  $\mathbf{Q}$  be a  $d$ -by- $d$  nonsingular matrix. Then, the multiplication of  $\tilde{\mathbf{A}}$  by  $\mathbf{Q}$  yields  $\mathbf{Q}\tilde{\mathbf{A}} = \mathbf{Q}\mathbf{F}\mathbf{W} + \mathbf{Q}\mathbf{N}$ . The assumption still remains valid for  $\mathbf{Q}\mathbf{F}$  due to the nonsingularity of  $\mathbf{Q}$ . Accordingly, if we can construct  $\mathbf{Q}$  so as to decrease the condition number of  $\mathbf{F}$ , the noise robustness of SPA may be improved by performing SPA on the input  $(\mathbf{Q}\tilde{\mathbf{A}}, r)$  instead of  $(\tilde{\mathbf{A}}, r)$ . In [8], Gillis and Vavasis proposed a procedure for constructing such a preconditioning matrix  $\mathbf{Q}$ . Here, we should pay attention to the fact that, even if  $\mathbf{Q}$  decreases the condition number of  $\mathbf{F}$ , the amount of noise could be expanded up to factor  $\|\mathbf{Q}\|$  since  $\|\mathbf{Q}\mathbf{N}\| \leq \|\mathbf{Q}\|\|\mathbf{N}\|$ .

### 2.2.1 Case of $d = r$

Now let us explain the procedure for constructing a preconditioning matrix in [8]. For simplicity, we will consider the noiseless case on  $\tilde{\mathbf{A}}$ . That is,  $\tilde{\mathbf{A}} = \mathbf{A}$ . We assume that  $\mathbf{A}$  satisfies Assumption 1, and in addition, assume that the dimension  $d$  and the factorization rank  $r$  coincide with each other. Under these assumptions,  $\mathbf{A}$  is an  $r$ -by- $m$  separable matrix and has an  $r$ -by- $r$  basis matrix  $\mathbf{F}$ . We set  $\mathcal{S} = \{\mathbf{a}_1, \dots, \mathbf{a}_m\}$  for the column vectors  $\mathbf{a}_1, \dots, \mathbf{a}_m$  of  $\mathbf{A}$ , and consider the optimization problem,

$$\begin{aligned} \mathbf{P}(\mathcal{S}) : \quad & \text{minimize} \quad -\log \det(\mathbf{L}), \\ & \text{subject to} \quad \mathbf{a}^\top \mathbf{L} \mathbf{a} \leq 1 \text{ for all } \mathbf{a} \in \mathcal{S}, \\ & \quad \mathbf{L} \succ \mathbf{0}. \end{aligned}$$

$\mathbf{L}$  is the decision variable.  $\mathbf{P}(\mathcal{S})$  corresponds to the formulation of computing the minimum volume enclosing ellipsoid (MVEE) centered at the origin for  $\mathcal{S}$ . It has been shown in [11, 8] that the optimal solution  $\mathbf{L}^*$  is given by  $(\mathbf{F}\mathbf{F}^\top)^{-1}$ . Therefore,  $(\mathbf{L}^*)^{1/2}$  can be used for the preconditioning matrix in order to improve the condition number of  $\mathbf{F}$ , since  $\kappa((\mathbf{L}^*)^{1/2}\mathbf{F})^2 = \kappa(\mathbf{F}^\top \mathbf{L}^* \mathbf{F}) = \kappa(\mathbf{I}) = 1$ . Next, we consider the noisy case on  $\tilde{\mathbf{A}}$ . Let  $\mathbf{L}^*$  be the optimal solution of  $\mathbf{P}(\mathcal{S})$  for  $\mathcal{S} = \{\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_m\}$  where  $\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_m$  are the column vectors of  $\tilde{\mathbf{A}}$ . In this case,  $\mathbf{L}^*$  does not completely match  $(\mathbf{F}\mathbf{F}^\top)^{-1}$ , but the difference between these two matrices is thought to be small if the amount of noise is also small. Therefore,  $(\mathbf{L}^*)^{1/2}$  could serve as a preconditioning matrix for restricting the condition number of  $\mathbf{F}$ .

We may need to add a further explanation of  $\mathbf{P}(\mathcal{S})$ . The origin-centered MVEE for the points in  $\{\pm \mathbf{a} : \mathbf{a} \in \mathcal{S}\}$  is given as  $\{\mathbf{x} \in \mathbb{R}^r : \mathbf{x}^\top \mathbf{L}^* \mathbf{x} \leq 1\}$  where  $\mathbf{L}^*$  is the optimal solution. The volume of the MVEE is  $c(r)/\sqrt{\det(\mathbf{L}^*)}$  where  $c(r)$  is the volume of a unit ball in  $\mathbb{R}^r$  and a real number depending on the dimension  $r$ . Since  $\text{rank}(\mathbf{A}) = r$  due to Assumption 1(a), the convex hull of the points in  $\{\pm \mathbf{a} : \mathbf{a} \in \mathcal{S}\}$  is full-dimensional in  $\mathbb{R}^r$ , and thus, the MVEE has a positive volume.  $\mathbf{P}(\mathcal{S})$  is a convex optimization problem. Efficient algorithms such as interior-point algorithms and the Frank-Wolfe algorithms have been developed and are now available for solving it; see, for instance, [10, 12] for the details on these algorithms.

Gillis and Vavasis showed in [8] that the preconditioner  $(\mathbf{L}^*)^{1/2}$  makes it possible to improve the noise robustness of SPA under Assumption 1 and  $d = r$ . Here is their result.

**Theorem 2** (Theorem 2.9 of [8]). *Let  $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{N}$  for  $\mathbf{A} \in \mathbb{R}^{d \times m}$  and  $\mathbf{N} \in \mathbb{R}^{d \times m}$ . Suppose that  $\mathbf{A}$  satisfies Assumption 1 and the condition  $d = r$ . Let  $\mathbf{L}^*$  be the optimal solution of  $\mathbf{P}(\mathcal{S})$  where  $\mathcal{S} = \{\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_m\}$  for  $\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_m$  of  $\tilde{\mathbf{A}}$ . If  $\mathbf{n}_i$  of  $\mathbf{N}$  satisfies  $\|\mathbf{n}_i\|_2 \leq \epsilon$  for all  $i = 1, \dots, m$  with*

$$\epsilon \leq O\left(\frac{\sigma_{\min}(\mathbf{F})}{r\sqrt{r}}\right),$$

*then, Algorithm 1 with the input  $((\mathbf{L}^*)^{1/2} \tilde{\mathbf{A}}, r)$  returns the output  $\mathcal{I}$  such that the size of the basis error of  $\mathcal{I}$  is up to  $O(\kappa(\mathbf{F})\epsilon)$ .*

The “size of the basis error” in the above statement should be clearly explained. Let  $\mathcal{I}$  be a subset of  $\{1, \dots, m\}$  with  $r$  elements, and suppose the elements are arranged in some order. Given a near-separable matrix  $\tilde{\mathbf{A}}$  of (3), the size of the basis error of  $\mathcal{I}$  is

$$\max_{j=1, \dots, r} \|\tilde{\mathbf{a}}_{\mathcal{I}(j)} - \mathbf{f}_j\|_2.$$

Gillis and Ma in [6] developed other types of preconditioning matrices for SPA and analyzed the noise robustness.

### 2.2.2 Case of $d \neq r$

The discussion in the previous section was made under Assumption 1 and the condition  $d = r$ . We shall consider the usual situation in which a near-separable matrix  $\tilde{\mathbf{A}}$  of (3) has  $d \neq r$ . The following approach was suggested in [8] for handling this situation. SVD plays a key role. SVD decomposes  $\tilde{\mathbf{A}}$  into  $\tilde{\mathbf{A}} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$  where  $\mathbf{U}$ ,  $\mathbf{V}$  and  $\mathbf{\Sigma}$  are the same as those of (2). By using  $\mathbf{\Sigma}$ , we construct a  $d$ -by- $m$  diagonal matrix  $\mathbf{\Sigma}^r$  such that

$$\mathbf{\Sigma}^r = \begin{cases} (\text{diag}(\sigma_1, \dots, \sigma_r, 0, \dots, 0), \mathbf{0}) & \text{if } d \leq m, \\ (\text{diag}(\sigma_1, \dots, \sigma_r, 0, \dots, 0); \mathbf{0}) & \text{otherwise.} \end{cases} \quad (4)$$

Let  $\mathbf{A}^r = \mathbf{U}\mathbf{\Sigma}^r\mathbf{V}^\top$ . This is the best rank- $r$  approximation matrix to  $\mathbf{A}$  under the matrix 2-norm or the Frobenius norm. Also, note that  $\tilde{\mathbf{A}} = \mathbf{A}^r$  holds if  $\tilde{\mathbf{A}}$  does not contain  $\mathbf{N}$ . We construct a matrix  $\mathbf{P} \in \mathbb{R}^{r \times m}$  such that

$$\begin{pmatrix} \mathbf{P} \\ \mathbf{0} \end{pmatrix} = \mathbf{U}^\top \mathbf{A}^r, \text{ equivalently, } \mathbf{P} = \text{diag}(\sigma_1, \dots, \sigma_r)(\mathbf{V}^r)^\top \quad (5)$$

where  $\mathbf{V}^r = (\mathbf{v}_1, \dots, \mathbf{v}_r) \in \mathbb{R}^{m \times r}$  for the column vectors  $\mathbf{v}_1, \dots, \mathbf{v}_r$  of  $\mathbf{V}$ . As we will see in Section 3.1,  $\mathbf{P}$  can be thought of as an  $r$ -by- $m$  near-separable matrix having an  $r$ -by- $r$  basis matrix. Therefore, we can apply the discussion in the previous section to  $\mathbf{P}$ . Let  $\mathcal{S} = \{\mathbf{p}_1, \dots, \mathbf{p}_m\}$  for  $\mathbf{P}$ . We compute the optimal solution  $\mathbf{L}^*$  of  $\text{P}(\mathcal{S})$  and run Algorithm 1 on  $((\mathbf{L}^*)^{1/2}\mathbf{P}, r)$ . Algorithm 2 summarizes each step of the preconditioned SPA. The description is almost the same as that of Algorithm 2 of [8].

---

#### Algorithm 2 Preconditioned SPA

---

**Input:** A  $d$ -by- $m$  real matrix  $\mathbf{A}$  and a positive integer  $r$ .

**Output:** An index set  $\mathcal{I}$ .

- 1: Compute the SVD of  $\mathbf{A}$ . Let  $\sigma_1, \dots, \sigma_r$  be the top  $r$  largest singular values, and  $\mathbf{v}_1, \dots, \mathbf{v}_r \in \mathbb{R}^m$  be the corresponding right singular vectors. Construct  $\mathbf{P} = \text{diag}(\sigma_1, \dots, \sigma_r)(\mathbf{V}^r)^\top \in \mathbb{R}^{r \times m}$  for  $\mathbf{V}^r = (\mathbf{v}_1, \dots, \mathbf{v}_r)$ .
  - 2: Compute the optimal solution  $\mathbf{L}^*$  of  $\text{P}(\mathcal{S})$  for  $\mathcal{S} = \{\mathbf{p}_1, \dots, \mathbf{p}_m\}$  where  $\mathbf{p}_1, \dots, \mathbf{p}_m$  are the column vectors of  $\mathbf{P}$ .
  - 3: Construct  $\mathbf{P}^\circ = (\mathbf{L}^*)^{1/2}\mathbf{P}$ . Run Algorithm 1 on the input  $(\mathbf{P}^\circ, r)$ , and output the index set  $\mathcal{I}$  obtained by the algorithm.
- 

### 2.3 Our Result and Its Comparison with Theorem 2 by Gillis and Vavasis

Gillis and Vavasis in [8] showed empirical results, suggesting that Algorithm 2 can improve the noise robustness of SPA. However, a formal analysis was not given. Here, we give it.

**Theorem 3.** Let  $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{N}$  for  $\mathbf{A} \in \mathbb{R}^{d \times m}$  and  $\mathbf{N} \in \mathbb{R}^{d \times m}$ . Suppose that  $r \geq 2$  and  $\mathbf{A}$  satisfies Assumption 1. If  $\mathbf{N}$  satisfies  $\|\mathbf{N}\|_2 = \epsilon$  with

$$\epsilon \leq \frac{\sigma_{\min}(\mathbf{F})}{1225\sqrt{r}},$$



then, Algorithm 2 with the input  $(\tilde{\mathbf{A}}, r)$  returns the output  $\mathcal{I}$  such that there is an order of the elements in  $\mathcal{I}$  satisfying

$$\|\tilde{\mathbf{a}}_{\mathcal{I}(j)} - \mathbf{f}_j\|_2 \leq (432\kappa(\mathbf{F}) + 4)\epsilon$$

for all  $j = 1, \dots, r$ .

The proof is given in Section 3. Note that in this paper we will use the notation  $\epsilon$  to describe the size of  $\|\mathbf{N}\|_2$  or  $\|\mathbf{n}_i\|_2$  for the noise matrix  $\mathbf{N}$  of a near-separable one. Let us compare our result, Theorem 3, with that of Gillis and Vavasis, Theorem 2. The main advantage of ours is in that it ensures the noise robustness of the preconditioned SPA for separable NMF problems without the condition  $d = r$ , while their result ensures it under that condition. The dimension  $d$  is usually greater than the factorization rank  $r$  in separable NMF problems derived from actual applications such as topic extraction from documents [4, 2, 11] and endmember detection in hyperspectral images [7, 8]. Therefore, our result can be used as a guide for seeing how robust the preconditioned SPA is to noise when handling such applications.

As we will see in Section 3,  $\mathbf{P}$  in step 1 of Algorithm 2 is an  $r$ -by- $m$  near-separable matrix having an  $r$ -by- $r$  basis matrix. Furthermore, Assumption 1 holds for the basis matrix of  $\mathbf{P}$ , if the amount of noise involved in an input matrix is small. Therefore, Theorem 2 can apply to  $\mathbf{P}$ , and this implies a similar result to ours.

**Proposition 1.** *Let  $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{N}$  for  $\mathbf{A} \in \mathbb{R}^{d \times m}$  and  $\mathbf{N} \in \mathbb{R}^{d \times m}$ . Suppose that Assumption 1 holds for  $\mathbf{A}$  in  $\tilde{\mathbf{A}}$ . If  $\mathbf{N}$  satisfies  $\|\mathbf{N}\|_2 = \epsilon$  with*

$$\epsilon \leq O\left(\frac{\sigma_{\min}(\mathbf{F})}{r\sqrt{r}}\right),$$

*then, Algorithm 2 with the input  $(\tilde{\mathbf{A}}, r)$  returns the output  $\mathcal{I}$  such that the size of the basis error of  $\mathcal{I}$  is up to  $O(\kappa(\mathbf{F})\epsilon)$ .*

The proof is given in Section 3. Although the size of the basis error in Proposition 1 is of the same order as ours, the allowable amount of noise is worse by a factor  $1/r$ .

Our allowed noise range is described using the norm of  $\mathbf{N}$ , while theirs is described by the norm of the column vectors  $\mathbf{n}_i$ . Hence, we shall rewrite our result in terms of the norm of  $\mathbf{n}_i$ . By taking account of the fact that  $\|\mathbf{N}\|_2 \leq \sqrt{m} \max_{i=1, \dots, m} \|\mathbf{n}_i\|_2$  for  $\mathbf{N} \in \mathbb{R}^{d \times m}$ , Theorem 3 implies the following corollary.

**Corollary 1.** *Let  $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{N}$  for  $\mathbf{A} \in \mathbb{R}^{d \times m}$  and  $\mathbf{N} \in \mathbb{R}^{d \times m}$ . Suppose the same conditions in Theorem 3 hold. If  $\mathbf{n}_i$  of  $\mathbf{N}$  satisfies  $\|\mathbf{n}_i\|_2 \leq \epsilon$  for all  $i = 1, \dots, m$  with*

$$\epsilon \leq O\left(\frac{\sigma_{\min}(\mathbf{F})}{\sqrt{rm}}\right),$$

*then, Algorithm 2 with the input  $(\tilde{\mathbf{A}}, r)$  returns the output  $\mathcal{I}$  such that the size of the basis error of  $\mathcal{I}$  is up to  $O(\kappa(\mathbf{F})\epsilon)$ .*

We see that  $1/\sqrt{m}$  emerges in the description of allowed noise range. When handling separable NMF problems from actual applications, it would be reasonable to suppose a situation where  $m$  corresponds to the number of data points, and it could be large. Therefore, the allowed noise range of our result becomes weaker in such a situation.



### 3 Analysis of the Noise Robustness of Algorithm 2

The main goal of this section is to prove Theorem 3. In the discussion of the proof, we will see that Theorem 2 implies Proposition 1.

#### 3.1 Preliminaries

Let  $\tilde{\mathbf{A}}$  be of the form given in (3). We shall consider Algorithm 2 on the input data  $(\tilde{\mathbf{A}}, r)$ . Step 1 computes the SVD of  $\tilde{\mathbf{A}}$ , and decomposes it into  $\tilde{\mathbf{A}} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$  where  $\mathbf{U}$ ,  $\mathbf{V}$  and  $\mathbf{\Sigma}$  are the same as those of (2). The rank- $r$  approximation matrix  $\tilde{\mathbf{A}}^r$  is given as  $\tilde{\mathbf{A}}^r = \mathbf{U}\mathbf{\Sigma}^r\mathbf{V}^\top$  by using  $\mathbf{\Sigma}^r$  of (4). We denote  $\tilde{\mathbf{A}} - \tilde{\mathbf{A}}^r$  by  $\tilde{\mathbf{A}}^{r,c}$ . Then,  $\tilde{\mathbf{A}}$  can be represented as

$$\tilde{\mathbf{A}} = \tilde{\mathbf{A}}^r + \tilde{\mathbf{A}}^{r,c} \quad (6)$$

by using  $\tilde{\mathbf{A}}^r$  and  $\tilde{\mathbf{A}}^{r,c}$  such that

$$\tilde{\mathbf{A}}^r = \mathbf{U}\mathbf{\Sigma}^r\mathbf{V}^\top \text{ and } \tilde{\mathbf{A}}^{r,c} = \mathbf{U}\mathbf{\Sigma}^{r,c}\mathbf{V}^\top.$$

Here, we let  $\mathbf{\Sigma}^{r,c} = \mathbf{\Sigma} - \mathbf{\Sigma}^r$ .

$\mathbf{P}$  in step 1 of Algorithm 2 is given as (5). Using relation (6), it can be rewritten as

$$\begin{aligned} (\mathbf{P}; \mathbf{0}) &= \mathbf{U}^\top \tilde{\mathbf{A}}^r \\ &= \mathbf{U}^\top (\tilde{\mathbf{A}} - \tilde{\mathbf{A}}^{r,c}) \\ &= \mathbf{U}^\top (\mathbf{A} + \mathbf{N} - \tilde{\mathbf{A}}^{r,c}) \\ &= \mathbf{U}^\top (\mathbf{A} + \bar{\mathbf{N}}) \\ &= \mathbf{U}^\top ((\mathbf{F}, \mathbf{F}\mathbf{K})\mathbf{\Pi} + \bar{\mathbf{N}}) \\ &= \mathbf{U}^\top (\mathbf{F} + \bar{\mathbf{N}}^{(1)}, \mathbf{F}\mathbf{K} + \bar{\mathbf{N}}^{(2)})\mathbf{\Pi} \\ &= \mathbf{U}^\top (\hat{\mathbf{F}}, \hat{\mathbf{F}}\mathbf{K} + \hat{\mathbf{N}})\mathbf{\Pi}. \end{aligned} \quad (7)$$

In the above, we have used the notation  $\bar{\mathbf{N}} \in \mathbb{R}^{d \times m}$ ,  $\bar{\mathbf{N}}^{(1)} \in \mathbb{R}^{d \times r}$ ,  $\bar{\mathbf{N}}^{(2)} \in \mathbb{R}^{d \times (m-r)}$ ,  $\hat{\mathbf{F}} \in \mathbb{R}^{d \times r}$ , and  $\hat{\mathbf{N}} \in \mathbb{R}^{d \times (m-r)}$  such that

$$\bar{\mathbf{N}} = \mathbf{N} - \tilde{\mathbf{A}}^{r,c}, \quad (8)$$

$$(\bar{\mathbf{N}}^{(1)}, \bar{\mathbf{N}}^{(2)}) = \bar{\mathbf{N}}\mathbf{\Pi}^{-1},$$

$$\hat{\mathbf{F}} = \mathbf{F} + \bar{\mathbf{N}}^{(1)}, \quad (9)$$

$$\hat{\mathbf{N}} = -\bar{\mathbf{N}}^{(1)}\mathbf{K} + \bar{\mathbf{N}}^{(2)}. \quad (10)$$

Accordingly,  $\mathbf{P}$  can be represented as

$$\mathbf{P} = (\mathbf{G}, \mathbf{G}\mathbf{K} + \mathbf{S})\mathbf{\Pi} \quad (11)$$

by using  $\mathbf{G} \in \mathbb{R}^{r \times r}$  and  $\mathbf{S} \in \mathbb{R}^{r \times (m-r)}$  such that

$$\mathbf{U}^\top \hat{\mathbf{F}} = \begin{pmatrix} \mathbf{G} \\ \mathbf{0} \end{pmatrix} \text{ and } \mathbf{U}^\top \hat{\mathbf{N}} = \begin{pmatrix} \mathbf{S} \\ \mathbf{0} \end{pmatrix}. \quad (12)$$

$\mathbf{P}^\circ$  in the step 3 of Algorithm 2 is

$$\mathbf{P}^\circ = (\mathbf{G}^\circ, \mathbf{G}^\circ\mathbf{K} + \mathbf{S}^\circ)\mathbf{\Pi}$$

by using  $\mathbf{G}^\circ \in \mathbb{R}^{r \times r}$  and  $\mathbf{S}^\circ \in \mathbb{R}^{r \times (m-r)}$  such that

$$\mathbf{G}^\circ = (\mathbf{L}^*)^{1/2} \mathbf{G} \text{ and } \mathbf{S}^\circ = (\mathbf{L}^*)^{1/2} \mathbf{S}.$$

Hence,  $\mathbf{P}$  and  $\mathbf{P}^\circ$  are near-separable matrices, and  $(\mathbf{G}, \mathbf{G}\mathbf{K})\mathbf{\Pi}$  and  $(\mathbf{G}^\circ, \mathbf{G}^\circ \mathbf{K})\mathbf{\Pi}$  correspond to the separable matrices. In particular,  $\mathbf{G}$  and  $\mathbf{G}^\circ$  are the basis matrices, and  $(\mathbf{0}, \mathbf{S})\mathbf{\Pi}$  and  $(\mathbf{0}, \mathbf{S}^\circ)\mathbf{\Pi}$  are the noise matrices. It should be noted that  $\mathbf{P}$  and  $\mathbf{P}^\circ$  are  $r$ -by- $m$  near-separable matrices and these have  $r$ -by- $r$  basis matrices.

### 3.2 Proof of Proposition 1

Here, we prove several lemmas that will be necessary for the subsequent discussion. Similar statements have already been proven in [11]. More precisely, Lemmas 2, 3 and 4 correspond to (a), (b) and (c) of Lemma 7 of that paper, and we have included them here to make the discussion self-contained. After that, we prove Proposition 1. The proof is obtained from Theorem 2 together with the following lemmas.

**Lemma 1.** *Let  $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{N} \in \mathbb{R}^{d \times m}$ . Then,  $|\sigma_i(\tilde{\mathbf{A}}) - \sigma_i(\mathbf{A})| \leq \|\mathbf{N}\|_2$  for each  $i = 1, \dots, t$  where  $t = \min(d, m)$ .*

*Proof.* See Corollary 8.6.2 of [9]. □

**Lemma 2** (Lemma 7(a) of [11]).  $\|\bar{\mathbf{N}}\|_2 \leq 2\|\mathbf{N}\|_2$ .

*Proof.*  $\|\bar{\mathbf{N}}\|_2 = \|\mathbf{N} - \tilde{\mathbf{A}}^{r,c}\|_2 \leq \|\mathbf{N}\|_2 + \|\tilde{\mathbf{A}}^{r,c}\|_2$  since  $\bar{\mathbf{N}}$  is of the form (8). Also,  $\|\tilde{\mathbf{A}}^{r,c}\|_2 = \sigma_{\max}(\tilde{\mathbf{A}}^{r,c}) = \sigma_{r+1}(\tilde{\mathbf{A}})$ . From Lemma 1 and  $\sigma_{r+1}(\mathbf{A}) = 0$ , we have  $\sigma_{r+1}(\tilde{\mathbf{A}}) \leq \|\mathbf{N}\|_2$ . Thus,  $\|\bar{\mathbf{N}}\|_2 \leq 2\|\mathbf{N}\|_2$ . □

**Lemma 3** (Lemma 7(b) of [11]). *Let  $\mathbf{s}_i$  be the column vector of  $\mathbf{S}$  in (11). Suppose that  $\mathbf{K}$  satisfies Assumption 1(b). Then,  $\|\mathbf{s}_i\|_2 \leq 4\|\mathbf{N}\|_2$  for each  $i = 1, \dots, m - r$ .*

*Proof.* We see from (12) and (10) that  $(\mathbf{s}_i; \mathbf{0}) = \mathbf{U}^\top \hat{\mathbf{n}}_i$  and  $\hat{\mathbf{n}}_i = -\bar{\mathbf{N}}^{(1)} \mathbf{k}_i + \bar{\mathbf{n}}_i^{(2)}$ . Here,  $\mathbf{k}_i$  and  $\bar{\mathbf{n}}_i^{(2)}$  are the column vectors of  $\mathbf{K}$  and  $\bar{\mathbf{N}}^{(2)}$ , respectively. Thus,  $\|\mathbf{s}_i\|_2 = \|-\bar{\mathbf{N}}^{(1)} \mathbf{k}_i + \bar{\mathbf{n}}_i^{(2)}\|_2 \leq \|\bar{\mathbf{N}}^{(1)}\|_2 \|\mathbf{k}_i\|_2 + \|\bar{\mathbf{n}}_i^{(2)}\|_2$ . From Lemma 2 and  $\|\mathbf{k}_i\|_2 \leq 1$  due to Assumption 1(b), we have  $\|\mathbf{s}_i\|_2 \leq 4\|\mathbf{N}\|_2$ . □

**Lemma 4** (Lemma 7(c) of [11]).  $|\sigma_j(\mathbf{F}) - \sigma_j(\mathbf{G})| \leq 2\|\mathbf{N}\|_2$  for each  $j = 1, \dots, r$ .

*Proof.* We see from (12) that  $\mathbf{G}$  and  $\hat{\mathbf{F}}$  has the relation  $(\mathbf{G}; \mathbf{0}) = \mathbf{U}^\top \hat{\mathbf{F}}$ . Since  $\mathbf{U}$  is an orthogonal matrix, the singular values of  $\mathbf{G}$  coincide with those of  $\hat{\mathbf{F}}$ . Also,  $\hat{\mathbf{F}}$  is of the form (9). Thus, from Lemma 1, we have  $|\sigma_j(\mathbf{F}) - \sigma_j(\mathbf{G})| = |\sigma_j(\mathbf{F}) - \sigma_j(\mathbf{F} + \bar{\mathbf{N}}^{(1)})| \leq \|\bar{\mathbf{N}}^{(1)}\|_2 \leq 2\|\mathbf{N}\|_2$ . The last inequality follows from Lemma 2. □

**Lemma 5.** *Let  $\tilde{\mathbf{a}}_k$  and  $\mathbf{p}_k$  be the column vectors of  $\tilde{\mathbf{A}}$  and  $\mathbf{P}$ . Also, let  $\mathbf{f}_j$  and  $\mathbf{g}_j$  be those of  $\mathbf{F}$  and  $\mathbf{G}$ . We have  $\|\tilde{\mathbf{a}}_k - \mathbf{f}_j\|_2 \leq \|\mathbf{p}_k - \mathbf{g}_j\|_2 + 3\|\mathbf{N}\|_2$  for any  $k$  and any  $j$  in  $\{1, \dots, r\}$ .*

*Proof.* We see from (7) and (6) that  $\tilde{\mathbf{a}}_k$  and  $\mathbf{p}_k$  are related as

$$\begin{pmatrix} \mathbf{p}_k \\ \mathbf{0} \end{pmatrix} = \mathbf{U}^\top \tilde{\mathbf{a}}_k^r \text{ and } \tilde{\mathbf{a}}_k = \tilde{\mathbf{a}}_k^r + \tilde{\mathbf{a}}_k^{r,c}$$

where  $\tilde{\mathbf{a}}_k^r$  and  $\tilde{\mathbf{a}}_k^{r,c}$  are the column vectors of  $\tilde{\mathbf{A}}^r$  and  $\tilde{\mathbf{A}}^{r,c}$ . Also, from (12) and (9),  $\mathbf{g}_j$  and  $\mathbf{f}_j$  are related as

$$\begin{pmatrix} \mathbf{g}_j \\ \mathbf{0} \end{pmatrix} = \mathbf{U}^\top \hat{\mathbf{f}}_j \text{ and } \hat{\mathbf{f}}_j = \mathbf{f}_j + \bar{\mathbf{n}}_j^{(1)}$$

where  $\hat{\mathbf{f}}_j$  and  $\bar{\mathbf{n}}_j^{(1)}$  are the column vectors of  $\hat{\mathbf{F}}$  and  $\bar{\mathbf{N}}^{(1)}$ . Therefore,

$$\begin{aligned} \|\tilde{\mathbf{a}}_k - \mathbf{f}_j\|_2 &= \|(\tilde{\mathbf{a}}_k^r + \tilde{\mathbf{a}}_k^{r,c}) - (\hat{\mathbf{f}}_j - \bar{\mathbf{n}}_j^{(1)})\|_2 \\ &= \|\mathbf{U}^\top (\tilde{\mathbf{a}}_k^r - \hat{\mathbf{f}}_j) + \mathbf{U}^\top (\tilde{\mathbf{a}}_k^{r,c} + \bar{\mathbf{n}}_j^{(1)})\|_2 \\ &\leq \|\mathbf{U}^\top (\tilde{\mathbf{a}}_k^r - \hat{\mathbf{f}}_j)\|_2 + \|\mathbf{U}^\top \tilde{\mathbf{a}}_k^{r,c}\|_2 + \|\mathbf{U}^\top \bar{\mathbf{n}}_j^{(1)}\|_2 \\ &= \|\mathbf{p}_k - \mathbf{g}_j\|_2 + \|\tilde{\mathbf{a}}_k^{r,c}\|_2 + \|\bar{\mathbf{n}}_j^{(1)}\|_2. \end{aligned}$$

$\mathbf{U}$  is a  $d$ -by- $d$  orthogonal matrix in (7) that consists of the left singular vectors of  $\tilde{\mathbf{A}}$ . By Lemma 2, we can put an upper bound on the norm of  $\bar{\mathbf{n}}_j^{(1)}$  such that  $\|\bar{\mathbf{n}}_j^{(1)}\|_2 \leq \|\bar{\mathbf{N}}\|_2 \leq 2\|\mathbf{N}\|_2$ . Also, we can put an upper bound on the norm of  $\tilde{\mathbf{a}}_k^{r,c}$  such that  $\|\tilde{\mathbf{a}}_k^{r,c}\|_2 \leq \|\mathbf{N}\|_2$  due to  $\|\tilde{\mathbf{a}}_k^{r,c}\|_2 \leq \|\tilde{\mathbf{A}}^{r,c}\|_2 \leq \|\mathbf{N}\|_2$ . The last inequality is obtained in the same way as in the proof of Lemma 2. Therefore, we have  $\|\tilde{\mathbf{a}}_k - \mathbf{f}_j\|_2 \leq \|\mathbf{p}_k - \mathbf{g}_j\|_2 + 3\|\mathbf{N}\|_2$ .  $\square$

We are now ready to prove Proposition 1.

(*Proof of Proposion 1*). We show that Theorem 2 can apply to  $\mathbf{P}$  in step 1 of Algorithm 2, if the amount of noise  $\|\mathbf{N}\|_2$  is smaller than some level.  $\mathbf{P}$  can be written as (11), and hence, is an  $r$ -by- $m$  near-separable matrix having an  $r$ -by- $r$  basis matrix  $\mathbf{G}$ . We choose some real number  $\gamma$  such that  $\gamma > 2$ . Suppose that  $\|\mathbf{N}\|_2 \leq \frac{1}{\gamma}\sigma_{\min}(\mathbf{F})$ . It follows from Lemma 4 and Assumption 1(a) that  $\sigma_{\min}(\mathbf{G}) \geq \frac{\gamma-2}{\gamma}\sigma_{\min}(\mathbf{F}) > 0$ . Therefore, Assumption 1 holds for  $\mathbf{G}$ . This means that Theorem 2 can apply to  $\mathbf{P}$ . Its application leads to the following statement. Let  $\mathbf{L}^*$  be the optimal solution of  $\mathcal{P}(\mathcal{S})$  where  $\mathcal{S} = \{\mathbf{p}_1, \dots, \mathbf{p}_m\}$  for  $\mathbf{P}$ . If  $\mathbf{s}_i$  of  $\mathcal{S}$  satisfies  $\|\mathbf{s}_i\|_2 \leq \epsilon$  for all  $i = 1, \dots, m - r$  with  $\epsilon \leq O(\sigma_{\min}(\mathbf{G})/r\sqrt{r})$ , then, Algorithm 1 with  $((\mathbf{L}^*)^{1/2}\mathbf{P}, r)$  returns the output  $\mathcal{I}$  such that there is an order of the elements in  $\mathcal{I}$  satisfying  $\|\mathbf{p}_{\mathcal{I}(j)} - \mathbf{g}_j\|_2 \leq O(\kappa(\mathbf{G})\epsilon)$  for  $j = 1, \dots, r$ .

Suppose that

$$\|\mathbf{N}\|_2 \leq \frac{\sigma_{\min}(\mathbf{F})}{4\alpha r\sqrt{r} + 2} \quad (13)$$

for some  $\alpha \geq 1$ . Note also that  $\|\mathbf{N}\|_2 < \frac{1}{2}\sigma_{\min}(\mathbf{F})$ . From Lemmas 3 and 4,  $\|\mathbf{s}_i\|_2 \leq 4\|\mathbf{N}\|_2$  and  $\frac{\sigma_{\min}(\mathbf{F}) - 2\|\mathbf{N}\|_2}{\alpha r\sqrt{r}} \leq \frac{\sigma_{\min}(\mathbf{G})}{\alpha r\sqrt{r}}$ . Since  $\|\mathbf{N}\|_2$  is supposed to satisfy (13), we have

$$\|\mathbf{s}_i\|_2 \leq 4\|\mathbf{N}\|_2 \leq \frac{\sigma_{\min}(\mathbf{F}) - 2\|\mathbf{N}\|_2}{\alpha r\sqrt{r}} \leq \frac{\sigma_{\min}(\mathbf{G})}{\alpha r\sqrt{r}}.$$

This gives  $\|\mathbf{s}_i\|_2 \leq \frac{\sigma_{\min}(\mathbf{G})}{\alpha r \sqrt{r}}$ . Also, from Lemma 4,

$$\begin{aligned} \kappa(\mathbf{G}) &= \frac{\sigma_{\max}(\mathbf{G})}{\sigma_{\min}(\mathbf{G})} \\ &\leq \frac{\sigma_{\max}(\mathbf{F}) + 2\|\mathbf{N}\|_2}{\sigma_{\min}(\mathbf{F}) - 2\|\mathbf{N}\|_2} \\ &= \left(1 + \frac{1}{2\alpha r \sqrt{r}}\right) \kappa(\mathbf{F}) + \frac{1}{2\alpha r \sqrt{r}} \\ &\leq 2\kappa(\mathbf{F}) + 1. \end{aligned}$$

Therefore, from Lemmas 3 and 5, we have

$$\|\tilde{\mathbf{a}}_{\mathcal{I}(j)} - \mathbf{f}_j\|_2 - 3\|\mathbf{N}\|_2 \leq \|\mathbf{p}_{\mathcal{I}(j)} - \mathbf{g}_j\|_2 \leq \beta\kappa(\mathbf{G})\|\mathbf{s}_{i^*}\|_2 \leq (8\beta\kappa(\mathbf{F}) + 4\beta)\|\mathbf{N}\|_2$$

for some  $\beta > 0$  where let  $i^* = \arg \max_{i=1, \dots, m-r} \|\mathbf{s}_i\|_2$ . Consequently, if  $\|\mathbf{N}\|_2$  satisfies (13), then, Theorem 2 can apply to  $\mathbf{P}$ , and the application implies that  $\|\tilde{\mathbf{a}}_{\mathcal{I}(j)} - \mathbf{f}_j\|_2 \leq (8\beta\kappa(\mathbf{F}) + 4\beta + 3)\|\mathbf{N}\|_2$ .  $\square$

### 3.3 Proof of Theorem 3

The core part of the proof of Theorem 3 is to show that Theorem 1 can apply to  $\mathbf{P}^\circ$  if  $\|\mathbf{N}\|_2$  is small. To do this, we evaluate the upper bound on the condition number of  $\mathbf{G}^\circ$ . For the subsequent discussion, we need the following lemma.

**Lemma 6.** *Let  $\alpha$  be a real constant satisfying  $\alpha > \sqrt{2}$  and  $r$  be any integer satisfying  $r \geq 2$ . Suppose that  $\|\mathbf{N}\|_2 \leq \frac{\sigma_{\min}(\mathbf{F})}{\alpha\sqrt{r}}$  and  $\mathbf{K}$  satisfies Assumption 1(b). Then,*

- (a)  $\frac{\alpha\sqrt{r}-2}{\alpha\sqrt{r}}\sigma_{\min}(\mathbf{F}) \leq \sigma_{\min}(\mathbf{G}) \leq \sigma_{\max}(\mathbf{G}) \leq \sigma_{\max}(\mathbf{F}) + \frac{2}{\alpha\sqrt{r}}\sigma_{\min}(\mathbf{F})$ .
- (b)  $\|\mathbf{s}_i\|_2 \leq \frac{4\sigma_{\min}(\mathbf{G})}{\alpha\sqrt{r}-2}$  for  $i = 1, \dots, m-r$ .

*Proof.* Statement (a) follows from Lemma 4, and that (b) follows from Lemmas 3 and (a).  $\square$

#### 3.3.1 Upper Bound on the Condition Number of $\mathbf{G}^\circ$

We show that the condition number of  $\mathbf{G}^\circ$  is bounded from above by a real constant, if  $\|\mathbf{N}\|_2$  is smaller than some level. This can be proved by applying Theorem 2.8 in [8] to  $\mathbf{P}$  and by taking into account the discussion in the proof of Proposition 1. This implies that such an upper bound is obtained under  $\|\mathbf{N}\|_2 \leq O(\sigma_{\min}(\mathbf{F})/r\sqrt{r})$ . Meanwhile, we will provide a similar bound under  $\|\mathbf{N}\|_2 \leq O(\sigma_{\min}(\mathbf{F})/\sqrt{r})$ . This result can increase the allowable amount of noise  $\|\mathbf{N}\|_2$  by a factor  $1/r$  over that of the application of the theorem.

Our derivation follows that of [8]. More precisely, it will be shown through Lemmas 7 and 8 and Proposition 2, which have the following correspondence to the lemmas of that paper; Lemma 7 is Lemma 2.4, Lemma 8 is Lemmas 2.5 and 2.6, and Proposition 2 is Lemmas 2.6 and 2.7. Although a major part of our proof of the lemmas and proposition relies on the techniques developed in that paper, there are some differences. In particular, we use an alternate technique to prove

Proposition 2, and it allows us to derive the upper bound on the condition number of  $\mathbf{G}^\circ$  under  $\|\mathbf{N}\|_2 \leq O(\sigma_{\min}(\mathbf{F})/\sqrt{r})$ . Remark 2 details the differences between the proofs.

Now let us look at problem  $\mathbf{P}(\mathcal{S})$  in step 2 of Algorithm 2. Note that  $\mathcal{S} = \{\mathbf{p}_1, \dots, \mathbf{p}_m\}$  for the column vectors  $\mathbf{p}_1, \dots, \mathbf{p}_m$  of  $\mathbf{P}$ . For convenience, we will change the variable to  $\mathbf{P}(\mathcal{S})$  such that  $\mathbf{C} = \mathbf{G}^\top \mathbf{L} \mathbf{G}$  for a nonsingular matrix  $\mathbf{G}$  and consider the problem,

$$\begin{aligned} \mathbf{Q}(\mathcal{S}) : \quad & \text{minimize} \quad -\log \det(\mathbf{C}) + 2 \log \det(\mathbf{G}), \\ & \text{subject to} \quad \mathbf{p}^\top (\mathbf{G}^{-1})^\top \mathbf{C} \mathbf{G}^{-1} \mathbf{p} \leq 1 \text{ for all } \mathbf{p} \in \mathcal{S}, \\ & \quad \mathbf{C} \succ \mathbf{0}, \end{aligned}$$

which is equivalent to  $\mathbf{P}(\mathcal{S})$  under the nonsingular transformation  $\mathbf{G}$ .  $\mathbf{C}$  is the decision variable. Let  $\mathbf{C}^*$  be the optimal solution and  $\lambda_j$  denote the  $j$ th eigenvalue of  $\mathbf{C}^*$ . We will continue to use the notation  $\lambda_j$  for this purpose throughout this section. For the  $j$ th singular value  $\sigma_j$  of  $\mathbf{G}^\circ$ , we have  $\sigma_j = \lambda_j^{1/2}$  for  $j = 1, \dots, r$  since  $(\mathbf{G}^\circ)^\top \mathbf{G}^\circ = \mathbf{G}^\top \mathbf{L}^* \mathbf{G} = \mathbf{C}^*$ .

In Lemmas 7 and 8, we evaluate the lower and upper bounds on  $\det(\mathbf{C}^*)$  by using  $r$  and  $\lambda_j$ . These bounds give the inequality that  $r$  and  $\lambda_j$  need to satisfy. In Proposition 2, by using the inequality, we derive the lower and upper bounds on  $\lambda_j$  whose square root is equal to the singular value  $\sigma_j$  of  $\mathbf{G}^\circ$ . Lemma 7 is almost the same as Lemma 2.4 of [8].

**Lemma 7.** *Let  $\alpha$  be a real constant satisfying  $\alpha > \sqrt{2}$  and  $r$  be any integer satisfying  $r \geq 2$ . Suppose that  $\|\mathbf{N}\|_2 \leq \frac{\sigma_{\min}(\mathbf{F})}{\alpha\sqrt{r}}$  and  $\mathbf{K}$  satisfies Assumption 1(b). Then,*

$$\det(\mathbf{C}^*) \geq \left( \frac{\alpha\sqrt{r} - 2}{\alpha\sqrt{r} + 2} \right)^{2r}.$$

*Proof.* For an  $r$ -by- $r$  scaled identity matrix  $\theta \mathbf{I}$  with a positive real number  $\theta$ , we derive the upper bound on  $\theta$  such that  $\theta \mathbf{I}$  is feasible for  $\mathbf{Q}(\mathcal{S})$ . Since  $\mathbf{P}$  can be written as (11),  $\mathcal{S}$  contains two different types of vectors: one is  $\mathbf{g}_j$  and the other is  $\mathbf{G}\mathbf{k}_i + \mathbf{s}_i$  where  $\mathbf{g}_j$  is the column vector of  $\mathbf{G}$  and  $\mathbf{k}_i$  and  $\mathbf{s}_i$  are those of  $\mathbf{K}$  and  $\mathbf{S}$ . Therefore,  $\theta \mathbf{I}$  needs to satisfy two types of constraints,

$$\begin{aligned} \theta \mathbf{g}_j^\top (\mathbf{G}^{-1})^\top \mathbf{I} \mathbf{G}^{-1} \mathbf{g}_j &\leq 1 \text{ for } j = 1, \dots, r, \\ \theta (\mathbf{G}\mathbf{k}_i + \mathbf{s}_i)^\top (\mathbf{G}^{-1})^\top \mathbf{I} \mathbf{G}^{-1} (\mathbf{G}\mathbf{k}_i + \mathbf{s}_i) &\leq 1 \text{ for } i = 1, \dots, m - r. \end{aligned}$$

The first constraints hold if  $\theta \leq 1$ . For the second constraints, we have

$$\begin{aligned} (\mathbf{G}\mathbf{k}_i + \mathbf{s}_i)^\top (\mathbf{G}^{-1})^\top \mathbf{I} \mathbf{G}^{-1} (\mathbf{G}\mathbf{k}_i + \mathbf{s}_i) &= \|\mathbf{k}_i + \mathbf{G}^{-1} \mathbf{s}_i\|_2^2 \\ &\leq (\|\mathbf{k}_i\|_2 + \|\mathbf{G}^{-1}\|_2 \|\mathbf{s}_i\|_2)^2 \\ &\leq \left( 1 + \frac{4}{\alpha\sqrt{r} - 2} \right)^2 \\ &= \left( \frac{\alpha\sqrt{r} + 2}{\alpha\sqrt{r} - 2} \right)^2. \end{aligned}$$

The second inequality follows from Lemma 6 and also  $\|\mathbf{k}_i\|_2 \leq 1$  due to Assumption 1(b). Thus, the second constraints hold if  $\theta \leq \left( \frac{\alpha\sqrt{r} - 2}{\alpha\sqrt{r} + 2} \right)^2$ . Let  $\bar{\theta} = \left( \frac{\alpha\sqrt{r} - 2}{\alpha\sqrt{r} + 2} \right)^2$ . It satisfies  $0 < \bar{\theta} < 1$  because of  $\alpha\sqrt{r} > 2$ . Thus,  $\bar{\theta} \mathbf{I}$  is a feasible solution of  $\mathbf{Q}(\mathcal{S})$ . Accordingly, for the optimal solution  $\mathbf{C}^*$  and the feasible solution  $\bar{\theta} \mathbf{I}$ , we have

$$\det(\mathbf{C}^*) \geq \det(\bar{\theta} \mathbf{I}) = \bar{\theta}^r = \left( \frac{\alpha\sqrt{r} - 2}{\alpha\sqrt{r} + 2} \right)^{2r}.$$

□

Lemma 8 corresponds to Lemmas 2.5 and 2.6 of [8]. The lemmas of that paper need to put a condition on the amount of noise in order to derive of the upper bound on  $\det(\mathbf{C}^*)$ , while this lemma does not need to do so. This comes from the difference in the structures of near-separable matrices. Our lemma handles a near-separable matrix of the form (11). It has a preferable structure wherein the noise matrix contains an  $r$ -by- $r$  zero submatrix and this zero submatrix corresponds to a basis matrix.

**Lemma 8.** *Suppose that  $r \geq 2$ . Then,*

$$\det(\mathbf{C}^*) \leq \left( \frac{r - \lambda_j}{r - 1} \right)^{r-1} \lambda_j$$

for each  $j = 1, \dots, r$ .

*Proof.* We derive an upper bound on the sum of the eigenvalues of  $\mathbf{C}^*$  that is equivalent to  $\text{tr}(\mathbf{C}^*)$ . Since  $\mathbf{C}^*$  is feasible for  $\mathbf{Q}(\mathcal{S})$ , we have

$$\mathbf{g}_j^\top (\mathbf{G}^{-1})^\top \mathbf{C}^* \mathbf{G}^{-1} \mathbf{g}_j \leq 1, \text{ equivalently, } \|(\mathbf{C}^*)^{1/2} \mathbf{e}_j\|_2^2 \leq 1 \text{ for } j = 1, \dots, r.$$

Thus,

$$\begin{aligned} \lambda_1 + \dots + \lambda_r &= \text{tr}(\mathbf{C}^*) \\ &= \|(\mathbf{C}^*)^{1/2}\|_F^2 \\ &= \sum_{j=1}^r \|(\mathbf{C}^*)^{1/2} \mathbf{e}_j\|_2^2 \leq r. \end{aligned}$$

The arithmetic-geometric mean inequality means that  $(a_1 \times \dots \times a_r)^{1/r} \leq (a_1 + \dots + a_r)/r$  holds for nonnegative real numbers  $a_1, \dots, a_r$ . Therefore, we have, for each  $j = 1, \dots, r$ ,

$$\det(\mathbf{C}^*) = \lambda_1 \times \dots \times \lambda_r \leq \left( \frac{r - \lambda_j}{r - 1} \right)^{r-1} \lambda_j.$$

□

We denote

$$a = \left( \frac{\alpha\sqrt{2} - 2}{\alpha\sqrt{2} + 2} \right)^4. \quad (14)$$

The value of  $a$  is in  $0 < a < 1$  when  $\alpha > \sqrt{2}$ .

**Proposition 2.** *Let  $\alpha$  be a real constant satisfying  $\alpha > \sqrt{2}$  and  $r$  be any integer satisfying  $r \geq 2$ . Suppose that  $\|\mathbf{N}\|_2 \leq \frac{\sigma_{\min}(\mathbf{F})}{\alpha\sqrt{r}}$  and  $\mathbf{K}$  satisfies Assumption 1(b). Then,  $\lambda_j$  is bounded such that  $1 - \sqrt{1 - a} \leq \lambda_j \leq 1 + \sqrt{1 - a}$  for each  $j = 1, \dots, r$ .*

*Proof.* Lemmas 7 and 8 tells us that  $r$  and  $\lambda_j$  need to satisfy the inequality

$$\left( \frac{\alpha\sqrt{r}-2}{\alpha\sqrt{r}+2} \right)^{2r} \leq \left( \frac{r-\lambda_j}{r-1} \right)^{r-1} \lambda_j. \quad (15)$$

When  $r = 2$ , it becomes  $a \leq (2 - \lambda_j)\lambda_j$ . Thus, it is necessary for  $\lambda_j$  to satisfy  $1 - \sqrt{1-a} \leq \lambda_j \leq 1 + \sqrt{1-a}$ .  $\square$

**Remark 2.** Proposition 2 corresponds to Lemmas 2.6 and 2.7 of [8]. From the lower and upper bounds on  $\det(\mathbf{C}^*)$ , the lemmas of that paper construct an inequality that  $r$  and  $\lambda_j$  need to satisfy and determine the condition on  $\lambda_j$  such that the inequality holds for all  $r \geq 2$ . In contrast, this proposition only considers the case of  $r = 2$  for inequality (15), and determines the condition on  $\lambda_j$ .

Since we have  $\sigma_j = \lambda_j^{1/2}$  for the  $j$ th singular value  $\sigma_j$  of  $\mathbf{G}^\circ$ , this proposition gives the bounds on the singular values and condition number of  $\mathbf{G}^\circ$ .

**Corollary 2.** *Suppose that the same conditions in Proposition 2 hold. Then, we have*

$$\begin{aligned} \sigma_{\min}(\mathbf{G}^\circ) &\geq \left(1 - \sqrt{1-a}\right)^{1/2}, \\ \sigma_{\max}(\mathbf{G}^\circ) &\leq \left(1 + \sqrt{1-a}\right)^{1/2}, \\ \kappa(\mathbf{G}^\circ) &\leq \left(\frac{1 + \sqrt{1-a}}{1 - \sqrt{1-a}}\right)^{1/2}. \end{aligned}$$

### 3.3.2 Application of Theorem 1 to $\mathbf{P}^\circ$

As we saw in Section 3.1,  $\mathbf{P}^\circ$  in step 3 of Algorithm 2 is an  $r$ -by- $m$  near-separable matrix. In the proposition below, we show that Theorem 1 can apply to  $\mathbf{P}^\circ$ . Here, we should note that  $\mathbf{s}_i^\circ$ ,  $\mathbf{p}_{\mathcal{I}(j)}^\circ$  and  $\mathbf{g}_j^\circ$  in the proposition are the column vectors of  $\mathbf{S}^\circ$ ,  $\mathbf{P}^\circ$  and  $\mathbf{G}^\circ$ .

**Proposition 3.** *Let  $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{N}$  for  $\mathbf{A} \in \mathbb{R}^{d \times m}$  and  $\mathbf{N} \in \mathbb{R}^{d \times m}$ . Suppose that  $r \geq 2$  and  $\mathbf{A}$  satisfies Assumption 1. Let  $\epsilon$  be such that  $\|\mathbf{s}_i^\circ\|_2 \leq \epsilon$  for all  $i = 1, \dots, m - r$ . If*

$$\|\mathbf{N}\|_2 \leq \frac{\sigma_{\min}(\mathbf{F})}{\alpha\sqrt{r}}$$

*and  $\alpha = 1225$ , then, Algorithm 2 with the input  $(\tilde{\mathbf{A}}, r)$  returns the output  $\mathcal{I}$  such that there is an order of the elements in  $\mathcal{I}$  satisfying*

$$\|\mathbf{p}_{\mathcal{I}(j)}^\circ - \mathbf{g}_j^\circ\|_2 \leq (80\kappa(\mathbf{G}^\circ)^2 + 1)\epsilon$$

*for all  $j = 1, \dots, r$ .*

*Proof.* First, we show that a separable matrix  $\mathbf{G}^\circ(\mathbf{I}, \mathbf{K})\mathbf{\Pi}$  in  $\mathbf{P}^\circ$  satisfies Assumption 1. From Corollary 2, we have  $\sigma_{\min}(\mathbf{G}^\circ) \geq (1 - \sqrt{1-a})^{1/2} > 1$ . The last inequality strictly holds, since the



value of  $a$ , which is of (14), is in  $0 < a < 1$  due to  $\alpha = 1225$ . Thus, we see that (a) of Assumption 1 is satisfied. Furthermore, (b) of Assumption 1 is satisfied since the assumption is put on the  $\mathbf{K}$  of  $\tilde{\mathbf{A}}$ .

Next, we show that the size of the noise matrix  $(\mathbf{0}, \mathbf{S}^\circ)\mathbf{\Pi}$  in  $\mathbf{P}^\circ$  is within the range allowed by Theorem 1. Namely, we show that the inequality

$$\|\mathbf{s}_i^\circ\|_2 < \min\left(\frac{1}{2\sqrt{r-1}}, \frac{1}{4}\right) \frac{\sigma_{\min}(\mathbf{G}^\circ)}{1 + 80\kappa(\mathbf{G}^\circ)^2} \quad (16)$$

holds for  $i = 1, \dots, m-r$ . We derive the upper bound on the left-side value and the lower bound on the right-side value. The left-side value is bounded such that

$$\begin{aligned} \|\mathbf{s}_i^\circ\|_2 &= \|(\mathbf{L}^*)^{1/2} \mathbf{s}_i\|_2 \\ &= \|(\mathbf{L}^*)^{1/2} \mathbf{G} \mathbf{G}^{-1} \mathbf{s}_i\|_2 \\ &\leq \|(\mathbf{L}^*)^{1/2} \mathbf{G}\|_2 \|\mathbf{G}^{-1}\|_2 \|\mathbf{s}_i\|_2 \\ &\leq \frac{4(1 + \sqrt{1-a})^{1/2}}{\alpha\sqrt{r-2}}. \end{aligned}$$

The last inequality follows from Corollary 2 and Lemma 6. The right-side value is bounded such that

$$\begin{aligned} \min\left(\frac{1}{2\sqrt{r-1}}, \frac{1}{4}\right) \frac{\sigma_{\min}(\mathbf{G}^\circ)}{1 + 80\kappa(\mathbf{G}^\circ)^2} &\geq \min\left(\frac{1}{2\sqrt{r-1}}, \frac{1}{4}\right) \frac{(1 - \sqrt{1-a})^{3/2}}{81 + 79\sqrt{1-a}} \\ &> \frac{1}{81} \min\left(\frac{1}{2\sqrt{r-1}}, \frac{1}{4}\right) \frac{(1 - \sqrt{1-a})^{3/2}}{1 + \sqrt{1-a}}. \end{aligned}$$

The first inequality follows from Corollary 2. Therefore, the inequality

$$\begin{aligned} \frac{4(1 + \sqrt{1-a})^{1/2}}{\alpha\sqrt{r-2}} &< \frac{1}{81} \min\left(\frac{1}{2\sqrt{r-1}}, \frac{1}{4}\right) \frac{(1 - \sqrt{1-a})^{3/2}}{1 + \sqrt{1-a}} \\ \Leftrightarrow \quad 324b &< \min\left(\frac{1}{2\sqrt{r-1}}, \frac{1}{4}\right) (\alpha\sqrt{r-2}) \end{aligned} \quad (17)$$

implies that of (16). Here, we denote

$$b = \left(\frac{1 + \sqrt{1-a}}{1 - \sqrt{1-a}}\right)^{3/2}.$$

Note that the value of  $b$  is determined by  $\alpha$ , since  $a$  is given as (14).

We show that inequality (17) holds for any  $r \geq 2$  when  $\alpha = 1225$ . In the case of  $2 \leq r \leq 5$ , it is sufficient to show

$$324b < \frac{\sqrt{2}}{4} \alpha - \frac{1}{2}. \quad (18)$$

In the case of  $r \geq 6$ , the inequality becomes

$$324b < \frac{\alpha\sqrt{r-2}}{2\sqrt{r-1}}. \quad (19)$$

Let  $f(x) = \frac{\alpha\sqrt{x}-2}{2\sqrt{x-1}}$  for  $x \geq 2$ . The function  $f$  attains its minimum at  $x = \alpha^2/4$ . Thus, we can put a lower bound on the right-side value.

$$\begin{aligned} \frac{\alpha\sqrt{r}-2}{2\sqrt{r-1}} &\geq f(\alpha^2/4) \\ &= \frac{1}{2}\sqrt{\alpha^2-4} \\ &> \frac{1}{2}\alpha - \frac{1}{2}. \end{aligned}$$

Accordingly, we see that inequality (18) implies that of (19). For inequality (18) with  $\alpha = 1225$ , we have the relation

$$324b < 432.4 < 432.6 < \frac{\sqrt{2}}{4}\alpha - \frac{1}{2}.$$

Therefore, inequality (17) holds for any  $r \geq 2$  when  $\alpha = 1225$ . This leads us to conclude that inequality (16) holds for  $i = 1, \dots, m-r$ .

□

Let us remark on the choice of  $\alpha = 1225$  in the proposition.

**Remark 3.** The value of  $b$  in (18) is given by the composite function  $f$  in  $x$  such that  $f = f_2 \circ f_1$  with

$$f_1(x) = \left( \frac{\sqrt{2}x-2}{\sqrt{2}x+2} \right)^4 \text{ and } f_2(x) = \left( \frac{1+\sqrt{1-x}}{1-\sqrt{1-x}} \right)^{3/2}.$$

The function  $f$  is monotonically decreasing for  $x \geq 2$ , and the function value approaches to 1 as  $x$  goes to infinity. Thus,  $\alpha = 1225$  is the minimum integer that satisfies the inequality (18).

In Proposition 3, we evaluated the size of the basis error due to Algorithm 2 in terms of  $\mathbf{P}^\circ$ . The proof of Theorem 3 can be obtained by rewriting it as  $\tilde{\mathbf{A}}$  instead of  $\mathbf{P}^\circ$ .

(*Proof of Theorem 3*). The theorem supposes that  $r \geq 2$  and  $\mathbf{A}$  satisfies Assumption 1. Therefore, Proposition 3 tells us that, if  $\|\mathbf{N}\|_2 \leq \sigma_{\min}(\mathbf{F})/\alpha\sqrt{r}$  and  $\alpha = 1225$ , then, Algorithm 2 with the input  $(\tilde{\mathbf{A}}, r)$  returns the output  $\mathcal{I}$  such that there is an order of the elements in  $\mathcal{I}$  satisfying

$$\|\mathbf{p}_{\mathcal{I}(j)}^\circ - \mathbf{g}_j^\circ\|_2 \leq (80\kappa(\mathbf{G}^\circ)^2 + 1)\epsilon \quad (20)$$

for all  $j = 1, \dots, r$  where  $\epsilon$  satisfies

$$\|\mathbf{s}_i^\circ\|_2 \leq \epsilon. \quad (21)$$

The norm of  $\mathbf{s}_i^\circ$  is bounded from above by using  $\|\mathbf{N}\|_2$  such that

$$\begin{aligned} \|\mathbf{s}_i^\circ\|_2 &= \|(\mathbf{L}^*)^{1/2}\mathbf{s}_i\|_2 \\ &\leq \|(\mathbf{L}^*)^{1/2}\mathbf{G}\|_2\|\mathbf{G}^{-1}\|_2\|\mathbf{s}_i\|_2 \\ &\leq \frac{4(1+\sqrt{1-a})^{1/2}}{\sigma_{\min}(\mathbf{G})}\|\mathbf{N}\|_2. \end{aligned}$$

The last inequality follows from Corollary 2 and Lemma 3. We choose  $\epsilon$  as

$$\epsilon = \frac{4(1+\sqrt{1-a})^{1/2}}{\sigma_{\min}(\mathbf{G})}\|\mathbf{N}\|_2. \quad (22)$$

By this choice, the inequality (21) is satisfied for all  $i = 1, \dots, m - r$ . In what follows, we shall use  $k$  to denote  $\mathcal{I}(j)$  in (20) for simplicity. For the left-side of (20), we have

$$\begin{aligned}\|\mathbf{p}_k^\circ - \mathbf{g}_j^\circ\|_2 &= \|(\mathbf{L}^*)^{1/2}(\mathbf{p}_k - \mathbf{g}_j)\|_2 \\ &\geq \sigma_{\min}((\mathbf{L}^*)^{1/2})\|\mathbf{p}_k - \mathbf{g}_j\|_2 \\ &\geq \sigma_{\min}((\mathbf{L}^*)^{1/2}\mathbf{G})\sigma_{\min}(\mathbf{G}^{-1})\|\mathbf{p}_k - \mathbf{g}_j\|_2 \\ &\geq \frac{(1 - \sqrt{1-a})^{1/2}}{\sigma_{\max}(\mathbf{G})}\|\mathbf{p}_k - \mathbf{g}_j\|_2.\end{aligned}$$

The last inequality follows from Corollary 2. For the right side of (20), we have, from Corollary 2 and the choice of  $\epsilon$  such as (22),

$$(1 + 80\kappa(\mathbf{G}^\circ)^2)\epsilon \leq \left(1 + \frac{80(1 + \sqrt{1-a})}{1 - \sqrt{1-a}}\right) \frac{4(1 + \sqrt{1-a})^{1/2}}{\sigma_{\min}(\mathbf{G})}\|\mathbf{N}\|_2.$$

Accordingly,

$$\begin{aligned}\frac{(1 - \sqrt{1-a})^{1/2}}{\sigma_{\max}(\mathbf{G})}\|\mathbf{p}_k - \mathbf{g}_j\|_2 &\leq \left(1 + \frac{80(1 + \sqrt{1-a})}{1 - \sqrt{1-a}}\right) \frac{4(1 + \sqrt{1-a})^{1/2}}{\sigma_{\min}(\mathbf{G})}\|\mathbf{N}\|_2 \\ \Leftrightarrow \|\mathbf{p}_k - \mathbf{g}_j\|_2 &\leq \frac{4(81 + 79\sqrt{1-a})(1 + \sqrt{1-a})^{1/2}}{(1 - \sqrt{1-a})^{3/2}}\kappa(\mathbf{G})\|\mathbf{N}\|_2.\end{aligned}\tag{23}$$

By Lemma 6, the condition number of  $\mathbf{G}$  is bounded from above by using that of  $\mathbf{F}$  such that

$$\begin{aligned}\kappa(\mathbf{G}) = \frac{\sigma_{\max}(\mathbf{G})}{\sigma_{\min}(\mathbf{G})} &\leq \frac{\alpha\sqrt{r}}{\alpha\sqrt{r}-2}\kappa(\mathbf{F}) + \frac{2}{\alpha\sqrt{r}-2} \\ &= \frac{\alpha\sqrt{r}}{\alpha\sqrt{r}-2}(\kappa(\mathbf{F}) + 1) - 1.\end{aligned}$$

Here, we consider the function  $f(x) = \frac{\alpha\sqrt{x}}{\alpha\sqrt{x}-2}$  for  $x \geq 2$ . Since the function is monotonically decreasing for  $x \geq 2$ , we have

$$\kappa(\mathbf{G}) \leq \frac{\alpha\sqrt{2}}{\alpha\sqrt{2}-2}\kappa(\mathbf{F}) + \frac{2}{\alpha\sqrt{2}-2}.$$

By using this inequality, we replace  $\kappa(\mathbf{G})$  in (23) with  $\kappa(\mathbf{F})$ , and then

$$\|\mathbf{p}_k - \mathbf{g}_j\|_2 \leq \frac{4(81 + 79\sqrt{1-a})(1 + \sqrt{1-a})^{1/2}}{(1 - \sqrt{1-a})^{3/2}} \left( \frac{\alpha\sqrt{2}}{\alpha\sqrt{2}-2}\kappa(\mathbf{F}) + \frac{2}{\alpha\sqrt{2}-2} \right) \|\mathbf{N}\|_2.$$

Since  $\alpha = 1225$  and  $a$  is of the form (14), the above inequality implies that  $\|\mathbf{p}_k - \mathbf{g}_j\|_2 \leq (432\kappa(\mathbf{F}) + 1)\|\mathbf{N}\|_2$ . By Lemma 5, this inequality leads to  $\|\tilde{\mathbf{a}}_k - \mathbf{f}_j\|_2 \leq (432\kappa(\mathbf{F}) + 4)\|\mathbf{N}\|_2$ .  $\square$

## Acknowledgments

The author would like thank Nicolas Gillis of Université de Mons who provided comments and suggestions on the draft version of this paper. We also thank the anonymous referee for insightful comments that enhanced the quality of this paper significantly. This research was supported by the Japan Society for the Promotion of Science (JSPS KAKENHI Grant Number 15K20986, 26242027).

## References

- [1] U. M. C. Araújo, B. T. C. Saldanha, R. K. H. Galvão, T. Yoneyama, H. C. Chame, and V. Visani. The successive projections algorithm for variable selection in spectroscopic multi-component analysis. *Chemometrics and Intelligent Laboratory Systems*, 57(2):65–73, 2001.
- [2] S. Arora, R. Ge, Y. Halpern, D. Mimno, and A. Moitra. A practical algorithm for topic modeling with provable guarantees. In *Proceedings of the 30th International Conference on Machine Learning (ICML)*, 2013.
- [3] S. Arora, R. Ge, R. Kannan, and A. Moitra. Computing a nonnegative matrix factorization – Provably. In *Proceedings of the 44th symposium on Theory of Computing (STOC)*, pages 145–162, 2012.
- [4] S. Arora, R. Ge, and A. Moitra. Learning topic models – Going beyond SVD. In *Proceedings of the 2012 IEEE 53rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1–10, 2012.
- [5] P. Businger and G. H. Golub. Linear least squares solutions by Householder transformations. *Numerische Mathematik*, 7(3):269–276, 1965.
- [6] N. Gillis and W. -K. Ma. Enhancing pure-pixel identification performance via preconditioning. *SIAM Journal on Imaging Sciences*, 8(2):1161–1186, 2015.
- [7] N. Gillis and S. A. Vavasis. Fast and robust recursive algorithms for separable nonnegative matrix factorization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(4):698–714, 2014.
- [8] N. Gillis and S. A. Vavasis. Semidefinite programming based preconditioning for more robust near-separable nonnegative matrix factorization. *SIAM Journal on Optimization*, 25(1):677–698, 2015.
- [9] G. H. Golub and C. F. Van Loan. *Matrix Computation*. The Johns Hopkins University Press, 3rd edition, 1996.
- [10] L. G. Khachiyan. Rounding of polytopes in the real number model of computation. *Mathematics of Operations Research*, 21(2):307–320, 1996.
- [11] T. Mizutani. Ellipsoidal rounding for nonnegative matrix factorization under noisy separability. *Journal of Machine Learning Research*, 15:1011–1039, 2014.
- [12] P. Sun and R. M. Freund. Computation of minimum-volume covering ellipsoids. *Operations Research*, 52(5):690–706, 2004.
- [13] S. A. Vavasis. On the complexity of nonnegative matrix factorization. *SIAM Journal of Optimization*, 20(3):1364–1377, 2009.